



# Query keyword suggestion based on location proximity.

<sup>1</sup>Prachi Shivaji Sable, <sup>2</sup>Babruvan Ramrao Solunke

<sup>1</sup>Student, <sup>2</sup>Assistant Professor

<sup>1</sup>Department of Computer Science and Engineering,

<sup>1</sup>NBNSCOE, Solapur, India

**Abstract:** Keyword suggestion in web search avails users to access relevant data without having knowledge in accordance with precisely expressing their queries. If location of the user considered information will be more appropriate. The keyword suggestion technique helps to retrieve documents which relates to information provided by user and user's location. Search engine retrieves the documents, which are semantically relevant to the original query using KD graph and they have as result documents that correspond to objects near to the user's location. To grab both the semantic relevance between keyword queries and the distance between the resulting documents and the query issuer's location, a location-aware keyword query suggestion framework is designed & also a weighted keyword-document graph.

**Index Terms - Keyword queries, Spatial distance, Document proximity.**

## I. INTRODUCTION

Keyword suggestion (query suggestion) has become most fundamental features of commercial web search engines. After submitting a Keyword query, the user may not be satisfied with the outcomes, so the Keyword suggestion module of the search engine suggests a set of m keyword queries that clarifies the user's search in the right path. On click information from query logs are used for effective keyword suggestion methods. New Keyword suggestions can be fixed according to their semantic relevance to the original keyword query. The first challenge of Keyword Suggestion framework is how to adequately measure keyword query similarity while considering the spatial distance factor. In accordance to previous query suggestion approaches Keyword Suggestion designs and uses a keyword-document bipartite graph (KD graph), which connects the keyword queries with their relevant documents. Diverse to all previous approaches which ignore locations, Keyword Suggestion adjusts the weights on edges in the KD graph to capture not only the semantic relevance between keyword queries, but also the spatial distance between the document locations and the query issuer's location the popular graph distance measure is applied for Personalized Page Rank (PPR) a.k.a. random walk with restart (RWR) distance on the KD graph, starting from the user supplied query, to find to set of m keyword queries with the highest semantic relevance to user supplied query and spatial proximity to the user location.

## II. LITERATURE SURVEY

Following are some existing studies and approaches related for Query Keyword Suggestion Based on Location Proximity:-

In paper [1] author specifies about data space is one of the emerging approaches to managing heterogeneous data sources. However, data spaces differ from conventional data integration approaches. Heterogeneous data sources with a pay-per-use integration approach. When data spaces meet unstructured data, indexing is the only way to answer queries based on the best effort. As far as we know, we have not worked on the provision of relevant results for a query beyond those provided by the indexation without manual efforts.

A technique for mining a collection of user transactions with an internet search engine to discover clusters of similar queries and similar URLs which is author introduces In paper [2] To a search engine along with URLs which the user selected from among the candidates offered by the search engine the information is click through data each record consist of users query . By viewing this dataset as bipartite graph, with the vertices to identify related queries and URLs. One noteworthy feature of the algorithm is that it is content ignorant.

In paper [3] author describe Random Walk with Restart (RWR) is a popular measure to estimate the similarity between nodes and has been exploited in numerous applications. For tracking RWR scores on dynamic graphs in an efficient way has aroused much interest among data mining researchers with many real-world graphs are dynamic with frequent insertion/deletion of edges. However, for updating RWR in a generalized form. Most of models fail to guarantee exactness and convergence time

In paper [4] author discussed about a novel graph combination based rare query suggestion framework. The system divided into two major steps: 1.click the graph contains query-URL click information and the skip graph contains query-URL skip information construct two query- URL bipartite graphs from query logs, 2. perform random walk on each of the graphs, using the random walk with restart (RWR) technique.

In paper [10] author describes a novel bookmark-coloring algorithm (BCA) that computes authority weights over the web pages utilizing the web hyperlink structure. The Page-Rank vector defined for a page-specific teleportation is similar to the computed vector (BCV) is similar to. Meanwhile, BCA is very fast, and BCV is sparse. BCA also has important algebraic properties. Several BCVs can be leveraged

in computing arbitrary BCV via a straightforward algebraic process and hub BCVs can be efficiently computed and encoded, If several BCVs corresponding to a set of pages (called hub) are known.

### III. PROBLEM STATEMENT

To retrieve relevant documents close to the query issuer's location & design the Location-aware Keyword query Suggestion framework, for suggestions relevant to the user's information needs.

### IV. OBJECTIVES AND SCOPE

#### Objectives:-

- To search relevant document to the keyword using KD graph & current location aware suggestion using 'location aware edge weight adjustment' approach with help Partition-Based Algorithm.
- To show categories & images of suggested queries of nearby location using database, also show image or logo of particular thing on map using Google API.
- To conduct extensive experiments to evaluate the effectiveness and efficiency of Location aware Keyword Suggestion. The effectiveness of Location aware Keyword Suggestion framework compared to query suggestion that does not consider location.

#### Scope:-

- In the proposed system, user searches keyword suggestions to get location aware documents using LKS framework according to current location of user.

### V. COMPARATIVE ANALYSIS

#### System Model of Existing scheme:-

In existing scheme, system provides location-based query suggestions by simply selecting the users past search queries that have results close to the user's current location. These suggestions may be insufficient if the user did not perform any historical searches near user's current location.

In general, a query submitted to a search engine, suggests a list of related queries. Queries can be issued by the user to the search engine to tune or redirect the search process and the related queries are based on previously issued queries, and. Query clustering process is process the method proposed which is based on a in which groups of semantically similar queries are identified. In the query log of the search engine the clustering process uses the content of historical preferences of users registered. The working of method is discovers the related queries, and ranks them according to a relevance criterion.

For search engine it cluster similar queries to recommend URLs to frequently asked queries. Following are the four notions of query distance: (1) based on keywords or phrases of the query; (2) based on string matching of keywords; (3) based on common clicked URLs; and (4) based on the distance of the clicked documents in some pre-defined hierarchy.

#### System Model of Proposed scheme:-

In Proposed scheme, system enhance the keyword query suggestion using location proximity framework user fire a query which may be single word or phrase. Then from that input keywords are extracted using that keywords and documents i.e. set of geo-documents with these two factors keyword document graph is constructed. It is directed weighted bipartite graph. Edge weight adjustment is the next step is location aware. The location aware edge weight adjustment is done in two ways I) Keyword- Document edge weight adjustment and II) Document-Keyword edge weight adjustment. These edge weight adjustments are separately formed because of the database have set of document for a particular location, to retrieve location aware document from multiple documents to the keyword, two separate approaches used in location aware suggestion framework. The edge weight adjustment is done based on location of the query issuer and the node of the KD graph afterword's suggestion are recommended which depend on relevance of the keywords i.e. initial user need and closeness of the document.

#### System Architecture:-

Step 1: Admin have to upload the dataset. Then preprocessing on this dataset and store all information into database.

Step 2: User enters the query for searching the any document with location. Proposed system find the distance between the users enter query and document location. Find the distance of keyword and document.

Display Result: This distance display on map and recommend the place of the user.

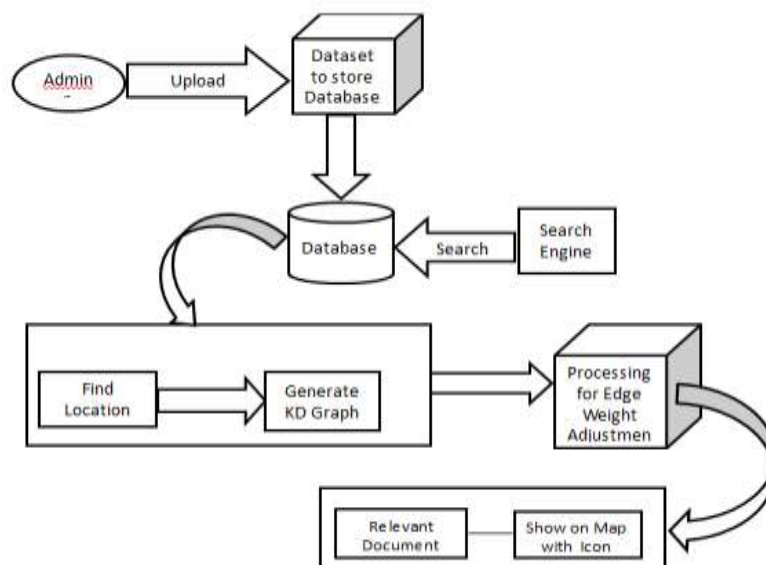


Figure 1: System Architecture

## VI. METHODOLOGY

Following figure shows the breakdown structure of the system

T1: Database: Database construction.

T2: Data Query Index: System will get query and provide index of query.

T3: KD Graph Construction: System will construct directed weighted bipartite graph.

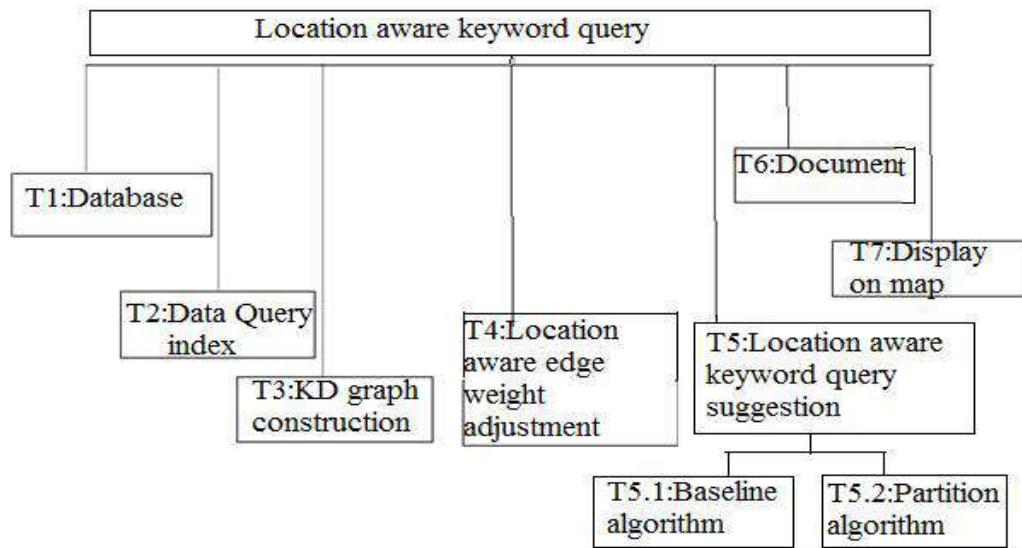


Figure 2. Breakdown Structure

T4: Location Aware Edge Weight adjustment: Conform the edge weights in the KD graph built on the spatial connections the middle of the area of the query backer and the hubs of the KD-graph.

T5: For Query Suggestion Location-Aware Keyword : Figure the random walk with restart procedure uses for every keyword queries a graph proximity score with admiration to kq, based on .

T5.1: Baseline Algorithm: For location-aware suggestions a baseline algorithm (BA) used .

T5.2: Partition Algorithm: A partition-based algorithm that isolates the keyword queries and the documents in the KD- graph G under aggregations.

T6: Documents: Matched documents display to the user.

T7: Display On Map: Matched location display on map with distance.

### Architecture Diagram:-

There were three modules admin login, User Query Search, Location Search, Recommended Place where in login we can create the user and log on to the particular user and In Query search we can provide our query it will produce result based on query. In location Search we can provide the location based on latitude and longitude results it will find nearest location based query and also recommend the nearest place.

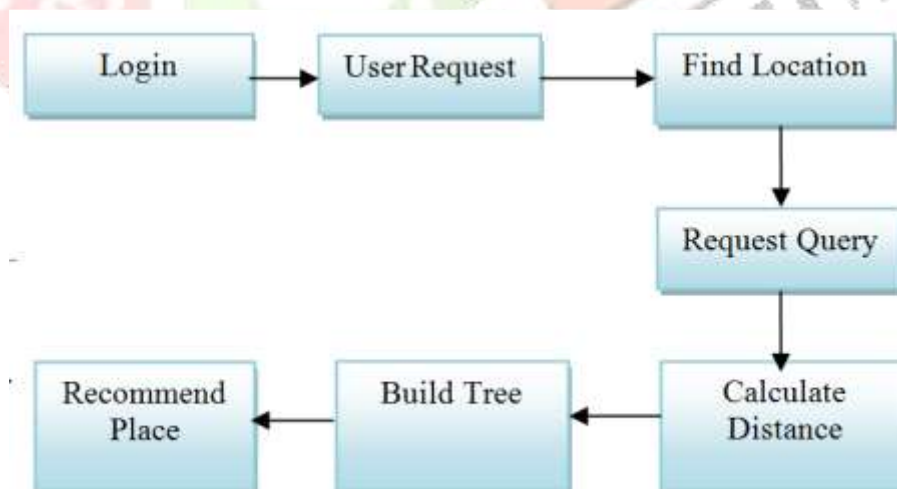


Figure3. Architecture Diagram

## VII. CONCLUSION

This study proposes relevant document search to the keyword using KD graph & current location aware suggestion using location aware edge weight adjustment with categories & images of suggested queries of nearby location, also show image or logo of particular thing on map by considering current location of user. System provides relevant user information need at the same time retrieve relevant document near to user location. For this evaluation user provide single keyword query then it calculate the distance based on query using partition algorithm.



### VIII. ACKNOWLEDGMENT

I would like to thanks Prof. B. R. Solunke for his immense interest, valuable guidance, constant inspiration and kind co-operation throughout the period of word undertaken, which has been instrumented in the success of our project. I also acknowledge our profound sense of gratitude to all the teachers who have been instrumental for providing me the technical knowledge and moral support to complete the project phase 1 with full understanding.

### REFERENCES

- [1] Z. Chen, H. T. Shen, X. Zhou, Y. Zheng, and X. Xie. Searching trajectories by locations: an efficiency study. In Proceedings of the 2010 ACM SIGMOD International Conference on Management of data, pages 255–266, 2010.
- [2] H.-P. Hsieh and C.-T. Li. Mining and planning time-aware routes from check-in data. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, pages 481–490, 2014.
- [3] W. T. Hsu, Y. T. Wen, L. Y. Wei, and W. C. Peng. Skyline travel routes: Exploring skyline for trip planning. In Mobile Data Management (MDM), 2014 IEEE 15th International Conference on, volume 2, pages 31–36, 2014.
- [4] Q. Yuan, G. Cong, and A. Sun. Graph-based point-of-interest recommendation with geographical and temporal influences. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, pages 659–668, 2014.
- [5] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. In Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, pages 325–334.
- [6] Y.-T. Wen, P.-R. Lei, W.-C. Peng, and X.-F. Zhou. Exploring social influence on location-based social networks. In Data Mining (ICDM), 2014 IEEE International Conference on, pages 1043–1048. IEEE, 2014.
- [7] H. Tong, C. Faloutsos, and J.-Y. Pan, “Fast random walk with restart and its applications,” in Proc. 6th Int. Conf. Data Mining, 2006, pp. 613–622.
- [8] Y. Fujiwara, M. Nakatsuji, M. Onizuka, and M. Kitsuregawa, “Fast and exact top-k search for random walk with restart,” Proc. VLDB Endowment, vol. 5, no. 5, pp. 442–453, Jan. 2012.
- [9] K. Avrachenkov, N. Litvak, D. Nemirovsky, E. Smirnova, and M. Sokol, “Quick detection of top-k personalized PageRank lists,” in Proc. 8th Int. Workshop Algorithms Models Web Graph, 2011, vol. 6732, pp. 50–61.
- [10] P. Berkhin, “Bookmark-coloring algorithm for personalized pagerank computing,” Internet Math., vol. 3, pp. 41–62, 2006.